

Notice of Allowability	Application No.	Applicant(s)
	10/767,151 Examiner	CHUNDI ET AL. Art Unit
	James H. Blackwell	2176

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTO-85) or other appropriate communication will be mailed in due course. THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS. This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. This communication is responsive to Amendment filed 09/18/2006.
2. The allowed claim(s) is/are 17-23 and 27-34.
3. Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All b) Some* c) None of the:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

* Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.
5. CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
 - (a) including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
 - 1) hereto or 2) to Paper No./Mail Date _____.
 - (b) including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).
6. DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. Notice of References Cited (PTO-892)
2. Notice of Draftsperson's Patent Drawing Review (PTO-948)
3. Information Disclosure Statements (PTO/SB/08),
Paper No./Mail Date _____
4. Examiner's Comment Regarding Requirement for Deposit
of Biological Material
5. Notice of Informal Patent Application
6. Interview Summary (PTO-413),
Paper No./Mail Date _____
7. Examiner's Amendment/Comment
8. Examiner's Statement of Reasons for Allowance
9. Other _____.

DETAILED ACTION

EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Steve Dicke on 12/19/2006. It is noted that with respect to Claim 37, the Applicant's representative did not choose one of several options suggested to overcome a 35 U.S.C. 101 rejection. The Examiner has chosen to cancel the claim as Applicant's representative was unavailable to make a timely response.

2. The Claims are amended as follows:

17. (Previously Presented) A method for clustering documents, including generating clusters with user perspective comprising:

receiving retrieval session logs;

performing log-based clustering on the session logs to generate session clusters;

representing each of the session clusters as a log-based document suitable for content based clustering;

receiving a plurality of documents that includes a first document that was accessed in one session and a second document that was not accessed in any of the sessions;

replacing the first document with one of the log-based documents, wherein said one of the log-based documents is associated with the session cluster that includes the first documents; and

performing content based clustering on at least one of the first document log-based document and the second document to generate clusters with user perspective.

18. (Previously Presented) The method of claim 17 wherein representing each of the session clusters as a log-based document suitable for content based clustering includes modifying each of the log-based documents referenced in the session cluster so that a Euclidean distance between the each of the log-based documents is the same.

19. (Previously Presented) The method of claim 17, wherein each of the session logs comprises a query used to retrieve documents.

20. (Previously Presented) The method of claim 17, wherein each of the session logs comprises a number of documents found to satisfy a query.

Art Unit: 2176

21. (Previously Presented) The method of claim 17, wherein each of the session logs comprises a list of documents opened by a user.

22. (Previously Presented) The method of claim 17, wherein each of the session logs comprises a length of time that a document was opened.

23. (Previously Presented) A method for clustering documents comprising:
generating a hybrid matrix of vectors comprising a first vector representing a first document and a second vector representing a log-based document cluster document;
and

clustering the documents using the hybrid matrix, wherein the hybrid matrix comprises:

accessing retrieval session logs;

clustering retrieval sessions into session clusters;

generating, a log-based document cluster for each session cluster by combining all documents opened during any retrieval session of the session cluster;

generating a log-based document cluster vector for each of the log-based document clusters;

replacing each document in the log-based document cluster with the log-based document cluster vector;

generating an individual document vector for each document not opened during any retrieval session; and

combining the log-based document cluster vector and the individual document cluster vector.

24-26 Cancelled

27. (Previously Presented) The method of claim 26 23 wherein the step of clustering retrieval sessions into session clusters comprises the steps of:

generating Boolean session vector for each retrieval session;
forming a matrix of the Boolean session vectors; and
applying a clustering algorithm to the matrix of the Boolean session vectors.

28. (Previously Presented) A system for clustering documents, the system comprising:

a storage for storing retrieval session logs; and
a processor connected to the storage, configured to cluster the retrieval sessions into session clusters,
generate, for each session cluster, a log-based document cluster,
generate a log-based document cluster vector for each of the log-based document cluster,
generate an individual document vector for each document not opened during any retrieval session,

cluster the documents using the log-based document cluster vectors and individual document vectors.

29. (Currently Amended) The system of claim 28 wherein the documents are stored in the storage.

30. (Currently Amended) The system of claim 28 further comprising:
a memory connected to the processor, for storage of a hybrid matrix comprising
the log-based document cluster vectors and the individual document
vectors.

31. (Previously Presented) A data processing system having session logs and documents, the system comprising:
a processor for executing program instructions; and
a media readable by the processor having a document clustering module having
a plurality of instructions, that when executed by the processor,
performs log-based clustering on the session logs to generate session
clusters,

converts the session clusters into a form suitable for content-based
clusters,

performs content-based clustering on the documents and session clusters in a form suitable for content-based clustering to generate document clusters with users' perspective.

32. (Currently Amended) The system of claim 31 wherein the document clustering module further comprises:

a session vector generation module for receiving the session logs and based thereon for generating a session vector for each session log;

a session cluster generation module coupled to the session vector generation module for receiving the session vectors and based thereon for generating session clusters;

a hybrid matrix builder for receiving the documents, coupled to the session cluster generation module, for receiving the session clusters and based thereon for generating a hybrid matrix having at least one log-based document; and

a topic generation module coupled to the hybrid matrix builder for receiving the hybrid matrix and based thereon for generating document clusters with users' perspective.

33. (Previously Presented) The system of claim 32 wherein the hybrid matrix builder further comprises:

a session document generation module for receiving session clusters and based thereon generates super documents; and

document modification module coupled to the session document generation module for receiving the super documents, for receiving the documents, and based thereon for generating the hybrid matrix.

34. (Previously Presented) The system of claim 31 wherein the media is one of a floppy disk, compact disc, a volatile memory, and a non-volatile memory.

35. Cancelled

Reasons For Allowance

3. The following is an examiner's statement of reasons for allowance:

The prior art of record, either alone or in combination, fails to teach, suggest, or render obvious the combined limitations of each of the independent claims.

In particular, the applicant's invention takes into account what pages on a given site (e.g., web site) users are actually viewing (and hence are interested in). It makes the assumption that within a given session (a user accessing pages on the web site sequentially over a determined time period), pages accessed are related (e.g., similar subject matter) since they were viewed consecutively within a finite time period. This analysis of user activity is performed on retrieval session logs. The results of this analysis are then included in content-based clustering of those results with the rest of the documents on the site, which were not accessed. The result of this content-based clustering analysis then produces results that give an administrator (e.g., web master)

Art Unit: 2176

an idea of what users were interested in over a set period of time, as well as help him determine future trends. The invention refers to this as "user perspective".

The prior art of record generally performs content-based clustering or log-based clustering independently making subsequent analysis of results for "user perspective" much more difficult and results may also vary wildly depending on the subsequent analysis technique.

4. Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Conclusion

5. Any inquiry concerning this communication or earlier communications from the examiner should be directed to James H. Blackwell whose telephone number is 571-272-4089. The examiner can normally be reached on Mon-Fri.
6. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.
7. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

James H. Blackwell
12/20/2006

H Herndon
Heather R. Herndon
Supervisory Patent Examiner
Technology Center 2100